

# **Análise de logs da plataforma Moodle utilizando técnicas de mineração de dados: um estudo de caso<sup>1</sup>**

**Jozemberg Gomes dos Santos Filho**

Curso de Licenciatura em Ciências da Computação – Universidade Federal do Paraíba  
(UFPB) – Campus IV – Litoral Norte  
58297-000 – Rio Tinto – PB– Brasil

jozemberg.filho@dce.ufpb.br

**Abstract.** Moodle is a virtual learning environment that offers many resources (forum, assignment, chat, quiz, etc ...) to create online courses. All executed actions by participants of the course (student, teacher or mediator) in the environment are stored in its database. In order to exploit this data and thus to extract some kind of implicit knowledge, this paper presents the results of an analysis of logs about the Introduction to Programming class, of the virtual Computing Course, at Federal University of Paraíba. For this, the steps of the Knowledge Discovery in Database process were followed, the KDD, where will be first made a selection and data preprocessing, soon after, will be selected required fields for application of Tableau Public tool generating infographics that will be interpreted and analyzed. One of the questions I had was regarding the day of posting and best times to the tutors are present in the environment, thus, it was possible to observe the students more use on the platform and what day and time more access the environment. It is expected that the results can help in decision making by the teacher and academic administration seeking to improve the teaching learning process.

**Resumo.** O Moodle é um ambiente virtual de aprendizagem que dispõe de diversos recursos (fórum, tarefa, chat, questionário, entre outros) para a criação de cursos online. Todas as ações realizadas pelos participantes do curso (aluno, professor ou tutor) no ambiente são registradas em sua base de dados. Com o intuito de explorar estes dados e, assim, extrair algum tipo de conhecimento implícito, este trabalho apresenta o resultado de uma análise de logs da disciplina de Introdução à programação do curso à distância de Licenciatura em Computação da Universidade Federal da Paraíba. Para isso, são seguidos os passos do processo de extração de conhecimento de dados, o KDD, onde será feito primeiramente uma seleção e pré-processamento dos dados, logo após, será selecionado campos necessários para aplicação da ferramenta Tableau Public gerando infográficos que serão interpretados e analisados. Um dos questionamentos que se tinha era em relação ao dia de postagem e melhores horários para os tutores estarem presentes no ambiente, sendo assim, foi possível observar o que os alunos mais utilizam na plataforma e qual o dia e horário que mais acessam o ambiente. A partir da montagem e análise de infográficos na ferramenta mencionada, espera-se que os resultados apresentados possam auxiliar a tomada de decisões por parte do docente e da gestão acadêmica buscando melhorar o processo de ensino-aprendizagem.

<sup>1</sup>Trabalho de Conclusão de Curso apresentado pelo aluno Jozemberg Gomes dos Santos Filho sob a orientação da professora Thereza Patrícia Pereira Padilha como parte dos requisitos para obtenção do grau de Licenciado em Ciência da Computação na UFPB Campus IV.

## 1. Introdução

O Brasil, sendo um país de grandes dimensões continentais, possui a difícil missão de proporcionar uma educação igualitária e de qualidade para o maior número de pessoas em todo território nacional. Visando tal proeza, o governo buscou a Educação a Distância (EaD) como um grande aliado, em que teve seu princípio com a utilização de cartas, posteriormente, com o advento da rádio difusão e à popularização dos televisores, os programas educativos eram realizados através de projetos do Ministério da Educação (MEC). Em seguida, os computadores e a Internet, juntamente com os Ambientes Virtuais de Aprendizagem (AVAs), aperfeiçoaram e ampliaram a metodologia da EaD no Brasil (DUARTE, 2002).

Após surgimento da Internet, no início de 1990, novos produtos e ferramentas foram desenvolvidos para explorar seus benefícios da melhor forma possível. Em meados da década 90, os AVAs apareceram com o objetivo de apoiar as atividades de ensino e aprendizagem através da Internet, oferecendo acesso flexível e novas formas de comunicação e avaliação para professores e alunos. Os AVAs permitem que professores criem recursos rapidamente sem a necessidade de desenvolver habilidades técnicas, pois é disponibilizado um conjunto integrado de ferramentas que possibilita uma melhor interação entre professor e aluno, inserção de diversos tipos de materiais, uma interface de fácil usabilidade e compreensão, dentre outras características. Segundo MERCADO (2002), o AVA é considerado o lugar onde professor e aluno estudam, debatem, discutem, constroem conhecimento, desenvolvem habilidades e atitudes, proporcionando um ambiente efetivo ao processo de ensino aprendizagem. A sala de aula convencional, por outro lado, deixa a desejar em alguns aspectos, caracterizando-se por:

Um espaço limitado em estruturas físicas, em um design convencional e homogêneo, em tempo e espaço. O espaço presencial é o espaço da “presença”, da territorialização, do apego ao aqui e agora, do encontro com o outro em uma mesma unidade de tempo e lugar, um lugar definido, com nome de país, estado, município, bairro, rua e número (SCHERER, 2005a, p. 47).

Alguns AVAs gratuitos e amplamente utilizados no Brasil são: Teleduc, AulaNet e o Moodle (*Modular Object Oriented-Dynamic Environment*). O Moodle e Teleduc possuem código fonte aberto, proporcionando sua disseminação e que diversas instituições o utilizem para pesquisar e produzir novas funcionalidades, tendo em vista um melhor acompanhamento do aluno no processo de ensino.

O Moodle está completamente integrado com o sistema acadêmico da UFPB, visando atender as necessidades da instituição. O administrador do Moodle tem permissão para ver todos os *logs* registrados no sistema inteiro, seja na homepage, seja nos cursos, seja nos recursos/atividades, seja no perfil dos usuários e acessível a ele pelo bloco de administração do site. Já os professores e tutores, geralmente (pois pode variar de uma instalação para outra), podem acessar *logs* apenas do curso, dos recursos/atividades e no perfil do usuário/alunos e disponível a eles pelo bloco de administração do curso.

A questão de acompanhamento dos *logs* é fundamental, pois é através dele que é possível emitir um parecer, zelando para que o aluno tenha sua participação de forma efetiva no curso. Ressalta-se que os relatórios disponibilizados pelo Moodle são do aspecto quantitativo e uma análise qualitativa dos dados de forma manual torna-se

inviável e trabalhosa, portanto, é necessário utilizar técnicas e ferramentas que facilitem a coleta, análise de informações em grandes base de dados para utilização de forma ágil e confiável. Diante deste contexto, foram feitos alguns questionamentos, como: Qual melhor dia da semana para postar atividades na plataforma? Qual recurso os alunos mais utilizam na plataforma? Qual melhor horário para os tutores auxiliar os alunos? Diante de tais questões, este trabalho tem como objetivo analisar e compreender os comportamentos dos alunos dentro da plataforma, onde será apresentado resultados de uma análise de *logs* do ambiente Moodle para a disciplina de Introdução à Programação, do curso de Licenciatura em Computação da UFPB Virtual, semestre 2016.1, utilizando a mineração de dados (MD).

Visando atender ao objetivo proposto, este artigo está estruturado da seguinte forma: na seção 2 serão apresentados a evolução da EaD no Ensino superior no Brasil e suas características, bem como o ambiente virtual de aprendizagem Moodle. Na seção 3 é exposto o processo de extração de conhecimento de bases de dados. Na seção 4 são apresentados os resultados obtidos em cada etapa do processo de extração de conhecimento de dados. Na seção 5 são mostradas as considerações finais deste trabalho.

## **2. Educação a Distância e Ambiente Moodle**

Se tratando de EaD é válido salientar às ações do governo federal acerca da formação de uma política para a EaD no Brasil. Vale destacar que, além da LDB 9.394/96, que inseriu a EaD ao sistema educacional brasileiro, várias portarias e decretos foram sendo publicados, buscando oferecer maior esclarecimento acerca dessa modalidade.

Em 2001, por meio da Portaria 2.253/01 foi aberta a possibilidade para que 20% da carga horária dos cursos presenciais pudessem ser realizadas a distância. Então, em 2004, a portaria 4.059/04, em vigor, passou a estabelecer que esses 20% referem-se ao total da carga horária de um curso, devendo a instituição organizar o currículo do curso de maneira que a somatória da carga horária do curso destinada as atividades.

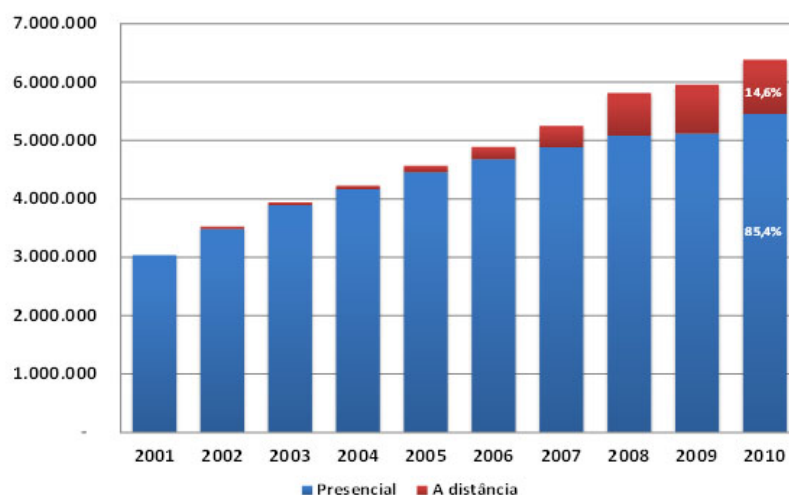
Para Llamas, a EaD é uma estratégia educativa baseada na aplicação da tecnologia à aprendizagem, e por isso, não obedece a limites de lugar, tempo, ocupação ou idade. Elementos que demandam novos papéis para alunos e professores, bem como novas atitudes e novos enfoques metodológicos. (ALVES et al. 2004).

Acompanhar a aprendizagem dos alunos na modalidade a distância não é uma tarefa trivial e sabe-se que a evasão é um dos maiores problemas em cursos ofertados a distância. Para o professor, a própria modalidade a distância é um dos fatores que dificulta acompanhar o processo de ensino aprendizagem em uma grande quantidade de alunos, e é fundamental que o docente faça o acompanhamento dos mesmos e tome decisões pedagógicas que os motivem para atingir seu objetivo no processo de ensino-aprendizagem para que o aluno não tenha frustração, desilusão e uma percepção negativa sobre a EAD. Cabe ressaltar que esse tipo de educação não se destina a qualquer tipo de aluno, a mudança do ambiente presencial para a modalidade a distância representa um desafio a mais, pois o aluno deve ser capaz de estudar de forma autônoma, sem a presença do professor como facilitador da aprendizagem.

No processo de estudo inicial, algumas dificuldades são encontradas, como: tempo, organização e planejamento sobre o horário de estudos, além de ter dificuldades em lidar com as tecnologias de informação e comunicação, podendo afetar o acompanhamento das atividades propostas pelo curso. A EAD baseada nas tecnologias de informação e comunicação requer uma alfabetização tecnológica que pode ser tornar um obstáculo insuperável tanto para alguns docentes, como os discentes.

Esta modalidade de ensino utilizava diversos meios, dentre eles: cartas, textos impressos, radiodifusão, entre outros. Alguns paradigmas foram quebrados com EaD, podemos destacar algumas: I – a necessidade de presença física do ensino frente a frente (educação presencial); II – o espaço físico delimitado para o contato entre aluno e professor; III – a estrutura organizacional da escola; IV – a turma de maneira homogênea; V – a forma de oratória para a disseminação do conhecimento; e IV – o tempo pré-fixado para a aprendizagem. (SILVA, 2004).

Segundo dados do Instituto Brasileiro de Geografia e Estatística (IBGE), no censo de 2000, 10,27% dos domicílios no Brasil possuíam computadores. Já em 2009, o número de domicílios com computadores já superava 34%, sendo 27,39% com acesso à internet IBGE (2004). O uso do computador junto com a internet foi possível devido à redução dos preços dos equipamentos. Já a internet teve seu custo reduzido após a entrada de diversas empresas no mercado, aumentando a concorrência e trazendo vários tipos de conexão (internet discada, banda larga, cabo, Wi-fi, rádio, 3G e etc) com custos menores. Assim, a EaD foi migrando da televisão para os meios virtuais. A Figura 1 apresenta o crescimento das matrículas na modalidade de ensino a distância no período de 2001 a 2010, atingindo 14,6% do total do número de matrículas.



**Figura 1 – Evolução no número de matrículas nas modalidades no Brasil 2001-2010.**  
**Fonte: SINEPES 2012.**

Com a democratização da Internet, foram criados vários tipos de software para dar suporte a educação. Dentre eles, destacam-se os Ambientes Virtuais de Aprendizagem, que permitem aos alunos e professores ter uma maior interação, disseminação do conteúdo, entre outras características. Para McKimm et al. (2003), um AVA consiste em um conjunto de ferramentas e recursos voltados ao processo de

ensino-aprendizagem que permite que um grande número de pessoas distribuídas geograficamente pelo mundo possa interagir em seus cursos.

Dentre os mais diversos AVAs existentes, destacam-se: AulaNet, TelEduc, Amadeus, Eureka, Moodle, e-Proinfo, Learning Space, WebCT, BlackBoard e SOLAR. O Moodle se sobressai pelo fato de ser um software livre, sendo uma alternativa imprescindível para se utilizar em quaisquer universidade pública ou privada. Como a educação normalmente se encontra com carência de recursos, o software livre é uma opção viável e que tem como preocupação principal suas funcionalidades (ALMEIDA, 2002). O Moodle é um AVA, desenvolvido por Dougiamas (2001), para funcionar como um sistema de gerência de curso. Diversas instituições de ensino superior que ofertam cursos na modalidade a distância utilizam o ambiente Moodle para promover a interatividade entre alunos, tutores e professores no processo de ensino e aprendizagem. A UFPB, por exemplo, utiliza este ambiente para a oferta dos seus cursos virtuais.

Destacam-se algumas vantagens da utilização do AVA Moodle, de acordo com Pulino Filho:

Muitos administradores de ambientes de aprendizagem têm declarado sua adesão ao Moodle principalmente em virtude de ser ele um sistema aberto, baseado em uma forte filosofia educacional, com uma comunidade de usuários crescente dia a dia que contribui para o desenvolvimento e apoio a novos usuários (PULINO FILHO, 2005, p. 5).

O sistema Moodle pode ser baixado gratuitamente da Internet. Uma das ferramentas disponíveis no Moodle são os relatórios, que são de aspecto quantitativo, servindo para questões estatísticas, onde auxiliam na gestão dos cursos ou disciplinas. Os relatórios disponibilizados pela plataforma são:

- **Relatório de Acesso (logs):** Gerenciamento de acesso de todos os participantes inscritos no curso, podendo ser de forma geral ou individual. Neste relatório é apresentado: dia e horário que o aluno acessou a plataforma, a ação realizada e IP de acesso.
- **Relatório de Logs Ativos:** Mostra os usuários logados no momento da consulta com os mesmos dados apresentados no relatório citado anteriormente.
- **Relatório de Atividades do Curso:** contém todas as atividades da disciplina ou curso, mostrando o número de visualizações de cada atividade e o último acesso da mesma.
- **Relatório de Participação do Curso:** Permite visualizar cada uma das atividades do curso e saber como e o quanto esta atividade foi acessada na disciplina.

### 3. Processo de Extração de Conhecimento de Bases de Dados

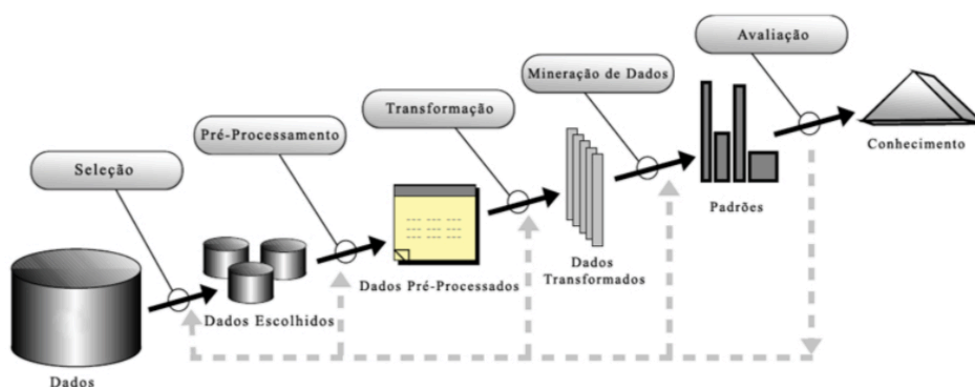
Atualmente, em todas as áreas, um considerável volume de dados tem sido sistematicamente coletado e armazenados. Segundo Fayyad et al. (1996), estes dados, produzidos e armazenados em larga escala, são inviáveis de serem lidos ou analisados por especialistas através de métodos tradicionais tais como planilhas de dados e relatórios informativos operacionais, em que o especialista testa sua hipótese contra a base de dados. Ou seja, as informações contidas nesses volumes de dados não estão bem

definidas explicitamente, uma vez que sendo dados operacionais não interessam quando estudados individualmente. Logo, é preciso transformar esses dados em informações, e, posteriormente, em conhecimento para que possa ser utilizado como suporte em situações de tomada de decisão.

O termo mineração de dados, por ser considerado multidisciplinar, as definições variam de acordo com o campo de atuação dos autores. Destacamos neste trabalho três áreas consideradas de maior relevância dentro da Mineração de dados: Estatística, Aprendizado de Máquina e Banco de Dados. Em Zhou (2003), é feita uma análise sobre essas três perspectivas:

- Em Hand et al. (2001): a definição é dada sobre uma perspectiva estatística: “Mineração de dados é a análise de grandes conjuntos de dados a fim de encontrar relacionamentos inesperados e de resumir os dados de uma forma que eles sejam tanto úteis quanto compreensíveis ao dono dos dados”.
- Em Cabena et al. (1998): a definição é dada sobre uma perspectiva de banco de dados: “Mineração de dados é um campo interdisciplinar que junta técnicas de máquinas de conhecimentos, reconhecimento de padrões, estatísticas, banco de dados e visualização, para conseguir extrair informações de grandes bases de dados”.
- Em Fayyad et al. (1996): a definição é dada sobre a perspectiva do aprendizado de máquina: “Mineração de Dados é um passo no processo de descoberta de conhecimento que consiste na realização da análise dos dados e na aplicação de algoritmos de descoberta que, sob certas limitações computacionais, produzem um conjunto de padrões de certos dados.”

Neste trabalho, a mineração de dados está voltada para a terceira perspectiva. Fayyad et al., 1996 propuseram um processo chamado de Extração de Conhecimento de Dados (*Knowledge Discovery in Databases - KDD*), que é iterativo e iterativo, dividido em cinco etapas, conforme pode ser visualizado na Figura 2.



**Figura 2 – Representação as etapas do processo KDD (Fayyad, 1996)**

Segue uma breve descrição de cada uma das etapas realizadas no processo KDD.

1. **Seleção:** obtenção dos dados a serem investigados na pesquisa.
2. **Pré-Processamento:** organização dos dados e tratamento das inconsistências encontradas quando necessário, como: remover ruídos, valores desconhecidos e dados duplicados. Assim, deve-se adotar estratégias de limpeza e de preparação para a obtenção de uma base de dados consistente.
3. **Transformação:** Após os dados serem selecionados e pré-processados, necessitam ser formatados e armazenados para que possam ser aplicados de forma adequada ao algoritmo de aprendizagem.
4. **Mineração de dados:** É a descoberta de conhecimento, onde os dados depois de transformados, serão lidos e interpretados aplicando técnicas para extração de conhecimento. Consiste na busca por padrões através da aplicação de algoritmos e técnicas computacionais específicas.
5. **Avaliação:** análise dos resultados e interpretação desses padrões através de visualizações, validando o conhecimento minerado em condições que podem ser facilmente compreendidas pelos indivíduos envolvidos no processo.

#### 4. Estudo de caso

Para a realização desse trabalho, foi escolhida a disciplina de Introdução à Programação do curso a distância de Licenciatura em Computação pela UFPB, semestre 2016.1. A escolha desta disciplina justifica-se pela sua importância no curso de Computação (componente curricular básico) e devido aos altos índices de retenção e evasão, conforme relatórios disponibilizados pela própria instituição. Nesta seção, são apresentados dados resultantes utilizando o *Tableu Public*, sendo uma ferramenta uma ferramenta gratuita e utilizada mundialmente e que suporta uma ampla gama de fontes de dados, a partir de arquivos locais, como CSV, Excel e dados armazenados em banco de dados. A mesma oferece de recursos interativos para análise e manipulação dos dados, onde é capaz de conectar-se a internet e disponibilizar suas aplicações finais através de um link para usuário.

A visualização de dados através de aspectos gráficos, estáticos ou interativos, destacando-se a importância da ligação entre imagem e texto com a finalidade de melhorar a comunicação em alguma coisa. A visualização de dados interativos permite a leitura das informações, possibilitando filtrar dados ao qual queremos visualizar, como também navegar em textos complementares. (PEDROZA et al. 2013)

Os dados resultantes da pesquisa serviram de fonte à geração de infográficos interativos, visando, assim, maior entendimento das informações que foram obtidas.

##### 4.1. Seleção, pré-processamento e transformação

Os dados foram coletados pela equipe de desenvolvimento da UFPB Virtual e entregues para análise no formato .CSV. Vale a pena ressaltar que a identificação dos envolvidos no curso foi totalmente excluída dos dados.

Analisando o arquivo dos dados manualmente, observou-se que existem 12 atributos, que são: *userid* (bigint), *eventname* (varchar), *component* (varchar), *action* (varchar), *target* (varchar), *objecttable* (varchar), *crud* (varchar), *edulevel* (tinyint), *contextlevel* (bigint), *timecreated* (bigint), *anonymous* (tinyint) e *origin* (varchar). Com relação à quantidade de registros, existem 12.913 instâncias registrando algum tipo de

interação dos envolvidos no curso na plataforma Moodle. A partir do *userid*, verificou-se ainda que esta turma (Introdução à Programação, semestre 2016.1) continha 60 alunos matriculados. Na Figura 3 é apresentada uma amostra dos dados obtidos.

	A	B	C	D	E	F	G	H	I	J	K
1	eventname,"component","action","target","objecttable","crud","edulevel","contextlevel","userid","anonymous","timecreated","origin"										
743	\mod_forum\event\course_module_viewed,"mod_forum","viewed","course_module","forum","r","2","70","25584","0","1456223921","web"										
744	\mod_forum\event\discussion_viewed,"mod_forum","viewed","discussion","forum_discussions","r","2","70","25584","0","1456223949","web"										
745	\mod_page\event\course_module_viewed,"mod_page","viewed","course_module","page","r","2","70","25584","0","1456223973","web"										
746	\mod_forum\event\course_module_viewed,"mod_forum","viewed","course_module","forum","r","2","70","23264","0","1456227027","web"										
747	\mod_forum\event\discussion_viewed,"mod_forum","viewed","discussion","forum_discussions","r","2","70","23264","0","1456227032","web"										
748	\mod_forum\event\discussion_viewed,"mod_forum","viewed","discussion","forum_discussions","r","2","70","27116","0","1456229676","web"										
749	\mod_forum\event\course_module_viewed,"mod_forum","viewed","course_module","forum","r","2","70","27116","0","1456229664","web"										
750	\mod_page\event\course module viewed,"mod page","viewed","course module","page","r","2","70","27116","0","1456229758","web"										

**Figura 3 – Amostra dos dados da tabela mdl\_logstore\_standard\_log**

Diante dos atributos existentes, observou-se ainda que os atributos *origin* (origem), *crud* (create, read, update, delete) e *anonymous* (anônimo) respectivamente continham apenas um valor possível (web), (r) e (0). Neste caso, estes atributos foram excluídos devido a não contribuir em nada para a análise das interações.

A tabela em estudo reflete ações realizadas pelo usuário a cerca da disciplina no ambiente virtual e é armazenada no campo *action*. Os valores considerados foram: *viewed*, *upload*, *sent*, *started*, *created*, *graded*, *update*, *submitted* e *reviewed*.

A fim de facilitar a legibilidade dos resultados, foi realizado uma transformação no formato dos atributos gerados a partir das datas, pois no banco de dados do Moodle são armazenadas no formato *Timestamp*, sendo assim, foi transformado para o formato de dias e horas, para visualizar de forma eficiente os campos gerados.

## 4.2. Mineração de dados e avaliação

Nesse momento da pesquisa, foram selecionados os campos de interesse para gerar infográficos interativos de forma automática através da ferramenta *Tableau Public* para que pudessemos adquirir representações com a finalidade de se explorar ou buscar novos conhecimentos. Foram escolhidos os campos respectivamente, dia da semana e hora, onde pode-se visualizar qual o dia da semana que os alunos mais utilizam o ambiente e seu horário de acesso.

Sabe-se que as frustrações dos alunos na EaD podem estar ligadas e motivadas por diversos fatores: instruções ambíguas no curso, problemas técnicos, meio de comunicação assíncrono, ausência de ajuda ou de resposta imediata por parte de tutores ou colegas, o que pode ser o elemento de evasão de um curso a distância. Dessa forma, através desses registros é possível identificar de maneira ágil o melhor horário e dia da semana que possam influenciar os professores e tutores na postagem de atividades e auxílio destas para se ter uma participação efetiva na plataforma, utilizando as ferramentas interativas que o ambiente disponibiliza para proporcionar uma melhor comunicação por ambas as partes, um feedback rápido dos exercícios enviados com a finalidade de motivar os alunos.

Sendo assim, como apresenta o infográfico (Figura 3), é possível visualizar que a maior quantidade de acesso na plataforma é a partir das 19:00 até as 23:00 horas do domingo. Nesse contexto, é mais viável postar atividades em dias que antecedem o maior fluxo de acesso, e assim, proporcionar que o maior número de alunos possam se beneficiar.



## DiasDaSemana x Horário de Acesso

Hora de	Dia Da Semana						
	segunda..	quarta-f..	quinta-f..	sexta-fei..	domingo	sabado	terça-fei..
0	6	4		19	5	6	7
1	24	2	12	3		8	9
2			5	14		6	5
3	2	11				11	
4	3	2			2		
5	15	1		3	5		3
6	39	27	18	61	31	26	42
7	51	28	62	55	97		37
8	90	62	53	155	119	26	45
9	97	166	118	86	153	16	45
10	162	204	149	90	125	84	72
11	105	108	86	98	118	20	63
12	56	63	44	100	77	46	85
13	113	61	61	158	87	46	29
14	100	76	59	71	132	86	76
15	106	123	87	16	151	72	129
16	89	98	107	18	135	115	171
17	72	77	143	28	144	93	90
18	85	85	54	41	164	128	126
19	196	91	61	53	550	54	110
20	248	103	95	78	614	72	134
21	99	90	105	74	479	128	125
22	210	77	59	51	351	170	32
23	101	21	68	29	228	78	23

**Figura 3 – Representação do dia da semana e hora que os alunos mais acessam a plataforma.**

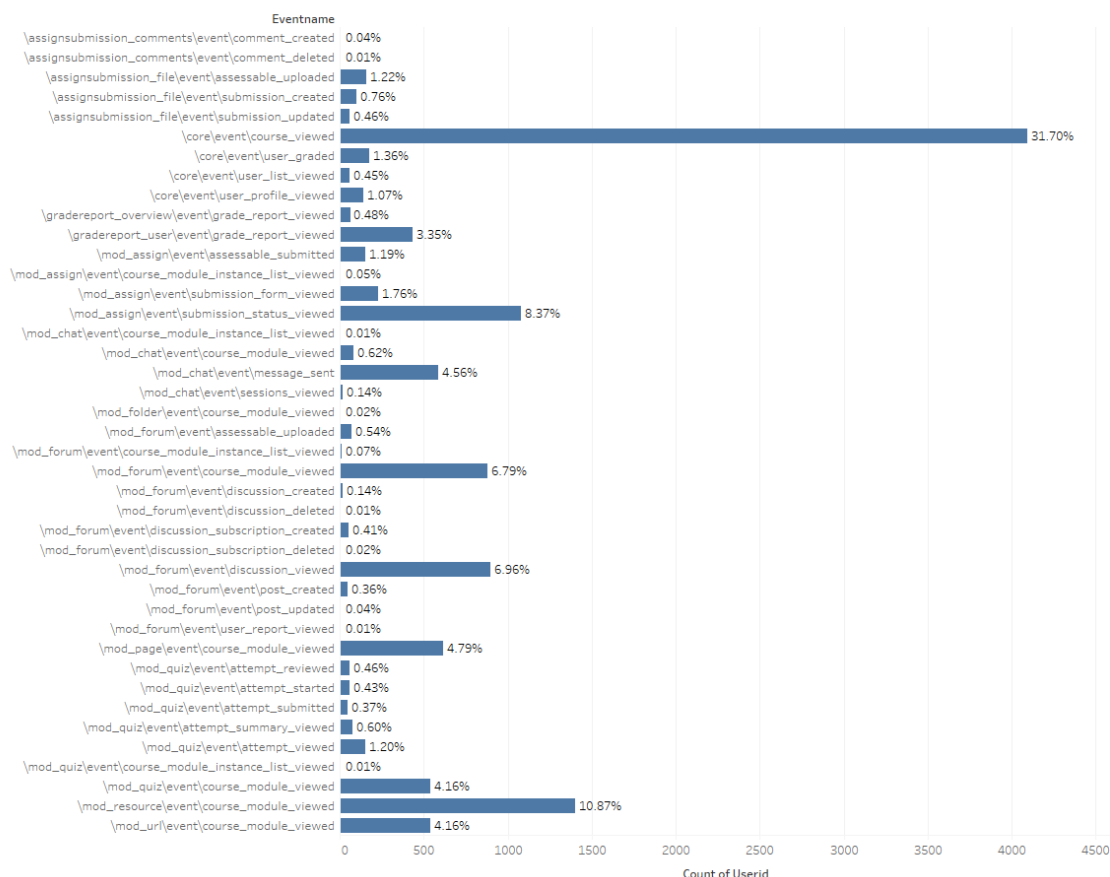
Outro fator que é bastante interessante para gestão acadêmica, é identificar na plataforma o comportamento oculto do alunos, como, quais as atividades mais utilizadas ou acessadas pelos mesmos. Com isso é possível auxiliar na tomada de decisões acerca de um melhor planejamento do curso, bem como as funcionalidades que irá ser mais utilizada pelos docentes, podendo contribuir e motivar uma maior intereção do aluno no curso.

O resultados do banco de dados analisado são apresentados a partir de uma visualização de dados utilizando o infográfico apresentando no figura 4. Podemos destacar as principais ações realizadas por os usuários, que são:

- 31,69% (4.093) visualizações dos alunos no curso;
- 10,87% (1.403) visualização de recursos;
- 8,37% (1.081) visualização do status da avaliação;
- 6,96% (899) visualização de discussão no fórum;

- 6,79% (877) visualização de post no fórum.

Descrição das ações mais utilizadas pelos usuários



**Figura 4 – Representação das atividades mais acessadas no ambiente.**

Como pode-se notar, as ações mais utilizadas na plataforma é a visualização dos módulos de curso, avaliação e fórum, podendo ou não comprometer o processo de aprendizagem, pois, talvez seja as ações que os gestores mais utilizam na plataforma. Através dos dados apresentados é possível verificar que se tem um índice baixo em relação a participação dos alunos nas seguintes questões: participação em chats, fórum e na criação de discussões e post.

## 5. Considerações finais

O banco de dados do ambiente *Moodle* possui aproximadamente 250 tabelas, e para extrair algumas informações foi feito um mapeamento da tabela de *logs*, assim, foi realizado um processo de refinamento para obter informações relevantes ao propósito desse estudo. Através da análise de dados obtida, foi possível identificar focos de possíveis melhorias, encontrando ações que possam ser executadas por parte da gestão acadêmica e pelos docentes através de tomada de decisões que influenciem no processo de ensino aprendizagem.

Através do processo de mapeamento dos campos de dia da semana e hora, foi possível identificar que os alunos poderiam ter um melhor aproveitamento de tempo para concluir atividades propostas pelo docente, e um possível intervalo de tempo e dias para que os tutores pudessem auxiliar os discentes nas possíveis dúvidas que venham aparecer. Foi observado também quais as ações mais utilizadas na plataforma, e como isso poderia influenciar no processo de ensino-aprendizagem.

. Acredita-se que a ferramenta *Tableau Public* possa contribuir no processo de ensino-aprendizagem na medida em que permite que os professores identifiquem de forma ágil e simplificada a necessidade de comunicação com os alunos, pois, como o software disponibiliza recursos de infografia, ressalta-se que quanto mais ampla for a possibilidade de manusear e manipular informações e retratar esses dados através de imagens, melhor a compreensão sobre tal representação.

Este trabalho, teve como objetivo minerar dados educacionais para auxiliar gestores acadêmicos a tornar o processo de ensino a distância mais eficaz para solucionar possíveis problemas educacionais. A utilização de técnicas de mineração de dados, e consequentemente o processo KDD, tem a contribuir e muito para a EaD, como para as demais áreas. Através da análise da base de dados do Moodle, podemos visualizar de maneira mais clara os relacionamentos entre os dados gerados pelas atividades propostas na plataforma, e com isso, espera-se trazer melhorias significantes para professores, tutores e alunos proporcionando e influenciando a esta modalidade a distância, uma forma de gerir melhor seus espaços de aula e contribuir no processo de ensino-aprendizagem.

## 6. Referências

- Almeida, Rubens Queiroz. Por que usar software livre? Disponível: <[http://www.dicasl.com.br/arquivo/por\\_que\\_usar\\_software\\_livre.php#.V9VsDzsW8d](http://www.dicasl.com.br/arquivo/por_que_usar_software_livre.php#.V9VsDzsW8d)> Acesso: Agosto/2016.
- Alves, Rêmulo Maia; Zambalde, André Luiz; & Figueiredo, Cristhiane Xavier. Ensino a Distância. UFLA/FAEPE. 2004.
- Alves. Lynn; Barros. Daniela; e Okada. Alexandra. MOODLE: Estratégias Pedagógicas e Estudos de Caso. 1. ed. Salvador – BA: EDUNEB 2009. Disponível: <http://repositorioaberto.uab.pt/bitstream/10400.2/2563/3/Livro%20Moodle.pdf>. Acesso: Agosto/2016.
- BRASIL. Lei no 9.394/96, de 20 de dezembro de 1996. Lei de Diretrizes e Bases da Educação Nacional. **Diário Oficial da União**, Brasília, DF, 1996. Disponível: <http://portal.mec.gov.br/seed/arquivos/pdf/tvescola/leis/lein9394.pdf>. Acesso: Setembro/2016.
- BRASIL. Ministério da Educação. Decreto no 5.159/2004. Aprova a estrutura regimental e o quadro demonstrativo dos cargos em comissão e das funções gratificadas do Ministério da Educação e dá outras providências. Brasília, DF, 2004a. Disponível: <<http://www.mec.gov.br>>. Acesso: Agosto/2016
- Cabena, P; Hadjinian, P; Stadler, R; Jaapverhees; Zanasi, A. Discovering Data Mining: From Concept to Implementation. Prentice Hall, 1998.

- Dougiamas, Martin; Taylor, Peter C. Interpretive analysis of an internet- based course constructed using a new courseware tool called Moodle. 2002. Disponível: <<http://dougiamas.com/writing/herdsa2002/>>. Acesso: Setembro/2016.
- Duarte, Zalina M. C. Educação a Distância (EAD): Estudo de fatores críticos de sucesso na gestão de cursos da região metropolitana de Belo Horizonte na visão dos tutores. Disponível: [http://www.fumec.br/anexos/cursos/mestrado/dissertacoes/completa/zalina\\_maria.pdf](http://www.fumec.br/anexos/cursos/mestrado/dissertacoes/completa/zalina_maria.pdf). Acesso: Setembro/2016.
- Fayyad, U; Piatetsky-Shapiro, G; Smyth, P. From Data Mining to Knowledge Discovery in Databases. American Association for Artificial Intelligence, 1996. Disponível: <https://www.aaai.org/ojs/index.php/aimagazine/article/viewFile/1230/1131>. Acesso: Agosto/2016.
- Hand, D; Mannila, H; Smyth, P. Principles of Data Mining. MIT Press, 2001. Disponível: [ftp://gamma.sbin.org/pub/doc/books/Principles\\_of\\_Data\\_Mining.pdf](ftp://gamma.sbin.org/pub/doc/books/Principles_of_Data_Mining.pdf). Acesso: Setembro/2016.
- Instituto Brasileiro de Geografia e Estatística - IBGE. Disponível: <<http://www.ibge.gov.br/home/estatistica/populacao/condicaodevida/indicadoresminimos/sinteseindicais2004/default.shtm>>. Acesso: Agosto/2016.
- Mckimm, J; Jollie, C; Cantillon, P. ABC of learning and teaching – Web based learning. Acessado em: , disponível: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1125774/pdf/870.pdf>. Acesso: Setembro/2016.
- Mercado, Luís P. L. (Org.). Novas Tecnologias na Educação: reflexões sobre a prática. Maceió. UFAL, 2002.
- Pedroza, N. P. Bezerra, E. P, Nicolau, M., Ferramentas para Elaboração de Infográficos: Um Estudo de Caso, Anais do Encontro Virtual de Documentação em Software Livre e Congresso Internacional de Linguagem e Tecnologia Online, Junho/2013.
- Scherer, Suely. Uma estética possível para a educação bimodal: aprendizagem e comunicação em ambientes presenciais e virtuais. Uma experiência em estatística aplicada à educação..., São Paulo, 2005a. 241 f. Tese (Doutorado em Educação: currículo), Pontifícia Universidade Católica de São Paulo Disponível: <http://www.catolicasc.org.br/arquivosUpload/5387401951342101780.pdf>. Acesso: Setembro/2016.
- Silva, Antônio C. R. Educação a Distância e o seu grande desafio: o aluno como sujeito de sua própria aprendizagem. (2004) Disponível: <http://www.abed.org.br/congresso2004/por/htm/012-TC-A2.htm>. Acesso: Agosto/2016 .
- Zhou, Z.-H. Three perspectives of data mining. Artificial Intelligence Journal, p.139–146, 2003. Disponível: [http://ac.els-cdn.com/S0004370202003570/1-s2.0-S0004370202003570-main.pdf?\\_tid=a1779c6e-7858-11e6-b42e-00000aab0f27&acdnat=1473623454\\_0d1b3ed76cedc287abd71524ae615476](http://ac.els-cdn.com/S0004370202003570/1-s2.0-S0004370202003570-main.pdf?_tid=a1779c6e-7858-11e6-b42e-00000aab0f27&acdnat=1473623454_0d1b3ed76cedc287abd71524ae615476). Acesso: Setembro/2016.